Prof. Dr.-Ing. Martin Bossert
M.Sc. Cornelia Ott / M.Sc. Jiongyue Xing
Exercise sheet 4

Task 4.1 (Shannon-Fano Code)

Let U be a random variable with the following probability distribution:

u	u_1	u_2	u_3	u_4	u_5	u_6	u_7
$P_U(u)$	0.3	0.2	0.1	0.1	0.1	0.1	0.1

a) Construct a binary prefix-free Shannon-Fano Code for U.

b) Calculate $\mathrm{E}[W]$ for the code constructed in a) and give an estimate for the entropy $H(U)$ (without calculating $H(U)$).

Task 4.2 (Huffman Code)

a) Construct a binary *Huffman Code* for the random variable U from Task 4.1 and compare the average length $\mathrm{E}[W]$ to the one obtained for the Shannon-Fano code.

b) Construct a ternary Huffman Code for U and show that the average code length satisfies $H(U)/\log_2 3 \leq \mathrm{E}[W] \leq H(U)/\log_2 3 + 1$.

c) The *efficiency* ν of a D-ary prefix-free code for a random variable U is defined as follows:

$$\nu = \frac{H(U)}{\mathrm{E}[W]\log_2 D}.$$

Compare the efficiency of the binary Shannon-Fano code code with the one of the binary Huffman code.

d) Compare the efficiency of the binary and the ternary Huffman code.

Task 4.3 (Tunstall-Algorithm)

Given is a binary memoryless source which outputs a 0 with probabilty 0.2 and a 1 with probability 0.8. These symbols should be mapped to binary strings of length 3. How many times must the codetree be extended during one run of Tunstall-Algorithm? How many sequences does the Tunstall-parser create? Construct the codetree and give a mapping.

Task 4.4 (Lempel-Ziv-Algorithm)

a) Apply the Lempel-Ziv-Algorithm to encode the sequence 00100101000110.

b) Decode the sequence

$$(0,0),(0,1),(1,0),(2,1),(4,0),(1,1),(5,1)$$

Task 4.5 (Efficiency of different source coding schemes)

Given a binary memoryless source (BMS) with the probabilities $P_U(0) = 0.9$ and $P_U(1) = 0.1$. In the following, we will consider the efficiency of different source coding schemes:

a) **(Block to variable length)**

Construct an optimal binary prefix-free code for a block length of $L = 1, 2, 3$ source bits and calculate the efficiency ν for each L, where

$$\nu = \frac{L \cdot H(U)}{\mathrm{E}[W] \log_2 D}.$$

b) **(Variable length to block)**

Construct an optimal binary block-code with codeword length $N = 1, 2, 3$ and calculate the efficiency

$$\nu = \frac{\mathrm{E}[Y] \cdot H(U)}{N \log_2 D},$$

where $\mathrm{E}[Y]$ is the average length of the coded message.

c) **(Variable length to variable length)**

Construct an optimal prefix-free code for the *Tunstall message set* with $M = 2, 4, 8$ codewords and calculate the efficiency

$$\nu = \frac{\mathrm{E}[Y] \cdot H(U)}{\mathrm{E}[W] \log_2 D},$$

where $\mathrm{E}[Y]$ is the average length of the coded message and $\mathrm{E}[W]$ is the average codeword length.